Study of deforestation with satellite data

Paco LARDY–NUGUES, Pauline BOUDY, Lisa BOUTEILLER, Hugo FEIDT
1A

22/05/2023

Supervisors: Mohamed JOUNI, Matthieu MULLER

Contents

1 Introduction

Remote sensing is the study of satellite images in order to obtain quantitative information. It has become very useful in many areas, including environmental monitoring.
The *Amazon* Rainforest is a major environmental issue, it is indeed what is called the *Earth's lung*, so its preservation is necessary.

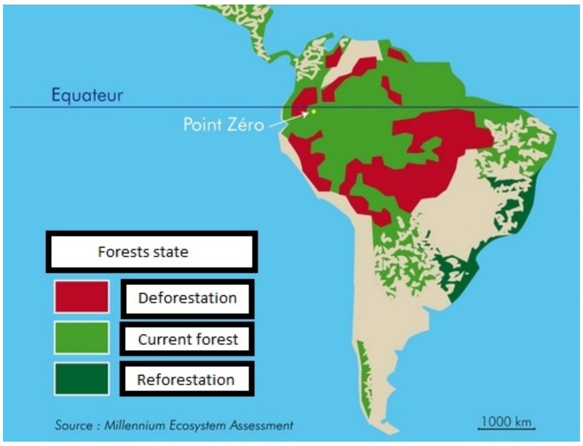


Figure 1.1 : State of the forests in the amazon [1]

Our goal here is to detect the progress of *deforestation* in the *Amazon* forest in recent years.
Our study will be carried out from the satellite *Sentinel 2* on different zones including crop and arid zones.

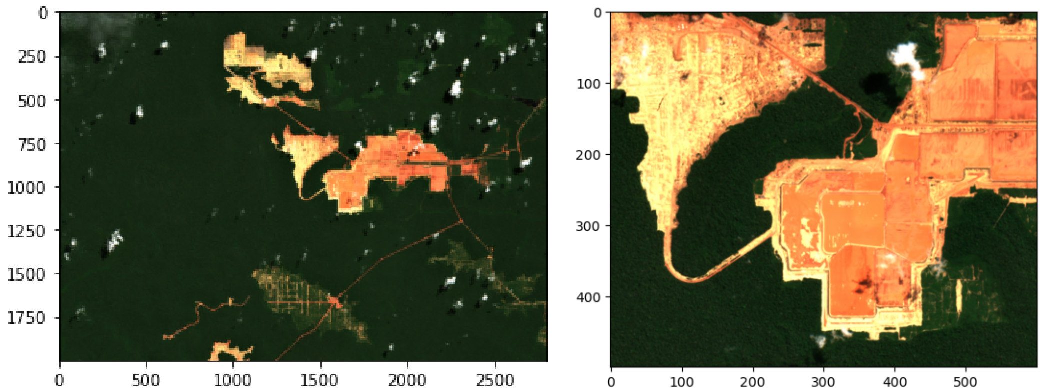


Figure 1.2 : Arid area in Amazon forest

However, the Sentinel 2 satellite was launched in 2016 encountered some issues which resulted in data loss for a year. To summarize, in some areas, we have little data, which is why we also retrieved images from *EarthExplorer*, an American satellite that has data that ranges from 2014 to 2022.

2 First Approach - Using the RGB Band

2.1 Definition of the Forest Rate and Thresholding

Detection of pathways, and other man-modified areas, can be detected from a thresholding using a vegetation index [2]:

$$NDVI_{green} = \frac{\rho_{green} - \rho_{rouge}}{\rho_{green} + \rho_{rouge}}.$$

We can roughly deforested areas:

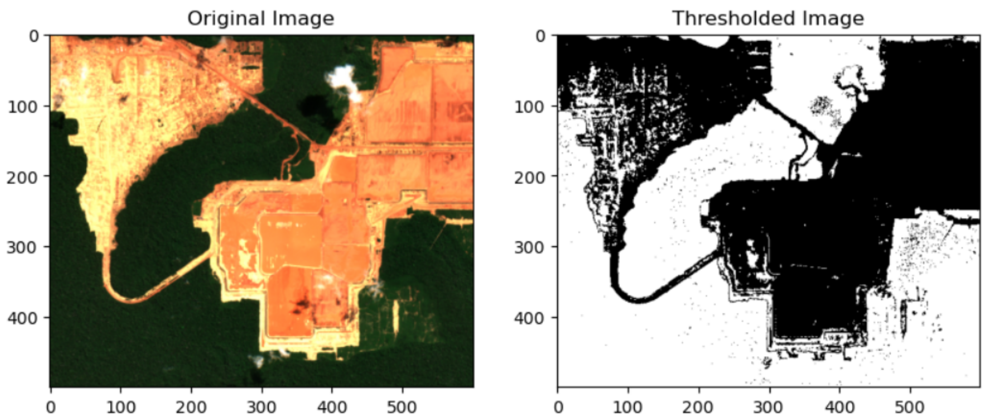


Figure 2.1.1 : Application of a Thresholding

And then, add an algorithm calculating the deforestation rate of an area defined by a pixel area:

If $\mathcal{A}(i,j)$ is the value of the pixel situated in (i,j), We can consider $\mathcal{K}$ the set of pixel taking a value above a certain threshold value.

$$\mathcal{K} = \left\{(i,j) \in \mathcal{P} \;\middle|\; \mathcal{A}(i,j) = \mathcal{R}\right\}.$$

Thus if

$$\mathcal{P} = N_{\text{horizontal pixel}} \times N_{\text{vertical pixel}}$$

Then

$$\mathcal{T}_{forest} = \frac{1}{\mathcal{P}} \sum_{i=0}^{N_{\mathrm{H}}-1} \sum_{j=0}^{N_{\mathrm{V}}-1} \mathbb{1}_{\mathcal{K}}(i,j).$$

We obtain for the above image , $\mathcal{T}_{forest} = 54.1\%$. We can therefore test on another zone:

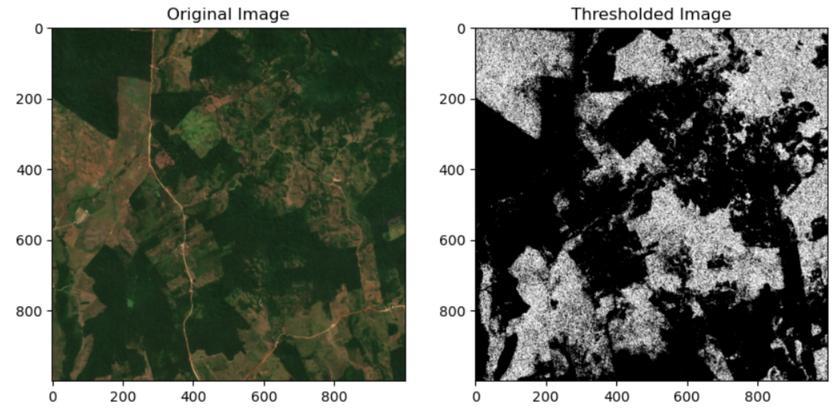


Figure 2.1.2 : $\mathcal{T}_{forest} = 27.5\%$

For this image, we had to change the formula of the *NDVI* to get better results. The colors of the crop fields are indeed quite different from the arid zones as seen previously:

$$NDVI_{blue} = \frac{\rho_{blue} - \rho_{rouge}}{\rho_{blue} + \rho_{rouge}}.$$

It is possible to be even more accurate by using another vegetation index. We notice that it is possible to accentuate what is not forest by using the vegetation index between the blue band and the green band:

$$NDVI_{green\&blue} = \frac{\rho_{blue} - \rho_{green}}{\rho_{blue} + \rho_{green}}.$$

Using the same method as before, we get the following images:

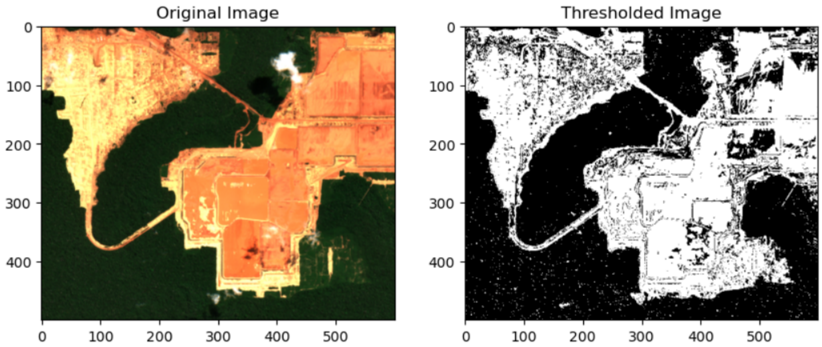


Figure 2.1.3 : $\mathcal{T}_{forest} = 54.8\%$

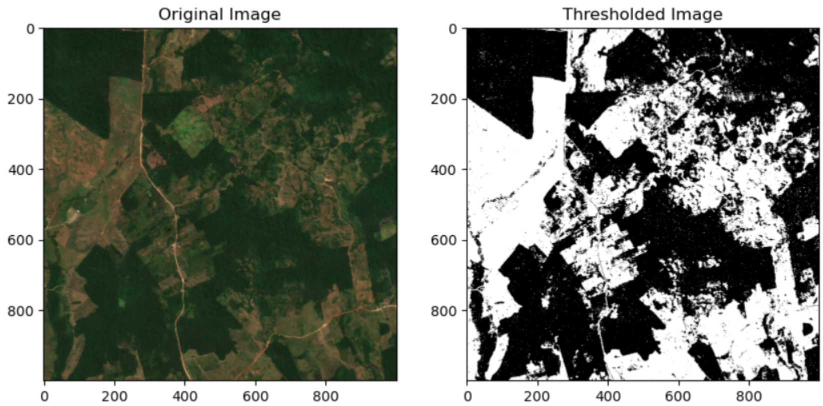


Figure 2.1.4: $\mathcal{T}_{forest} = 55.4\%$

We obtain something very different than the previous result for the *Figure 2.1.2*, more precise.

2.2 Application of a Thresholding on several years

We chose an area where we could have access to the exact same place for different years to avoid creating an algorithm to cut images to be in the same area. For each image, we are going to calculate the forest rate and plot all of this.

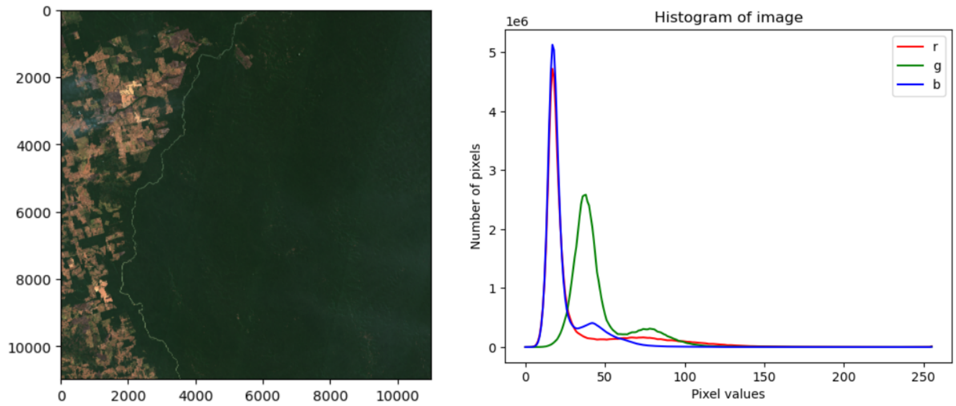
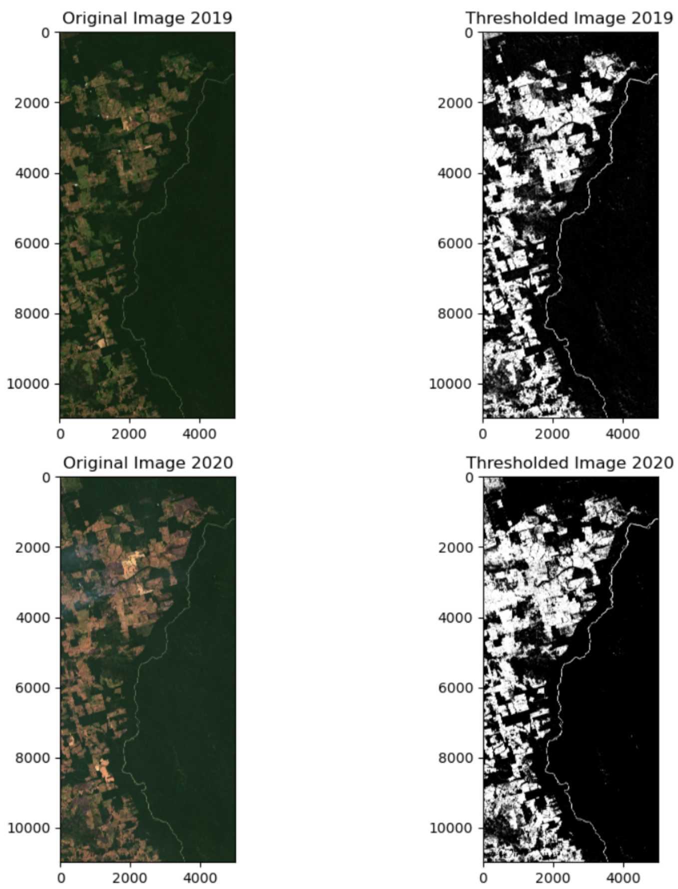


Figure 2.2.1 : Distribution histogram of 2019

We can see that it is interesting to use the $NDVI_{green\&blue}$ to apply a threshold (here with the value $\mathcal{R} = 3$ for 2019 and 2022, and then $\mathcal{R} = 2$ for the 2020's image due to a different contrast):

We can then display the *forest* evolution over 4 years:



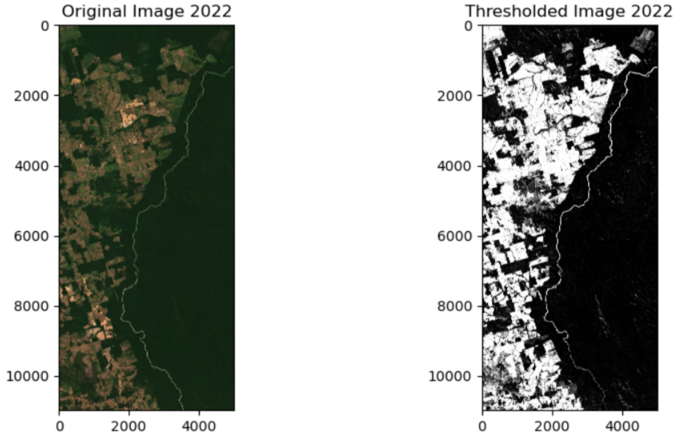


Figure 2.2.2 : Evolution of a forest in 4 years

And *plot* the *forest rate* over time. This leads to:

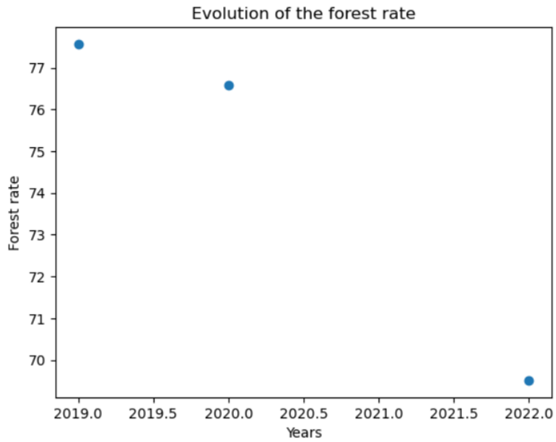


Figure 2.2.3 : Evolution of the forest rate

We were expecting a constant decrease. Furthermore, the second image (2020) had stronger contrast than the other two, which could have distorted the measurements. For these reasons, an other approach is required.

Unfortunately we don't have a lot of data and it is quite tough to see precisely the evolution, so we will download the data from the American EarthExplorer satellite, which contains data that ranges from 2014 to 2022 in another zone. To do this, we first applied a pre-processing to the images that contained a lot of clouds, by coloring the clouds the same color as the surroundings.

We also did a pre-processing to apply a contrast between 2019 and 2020 images:

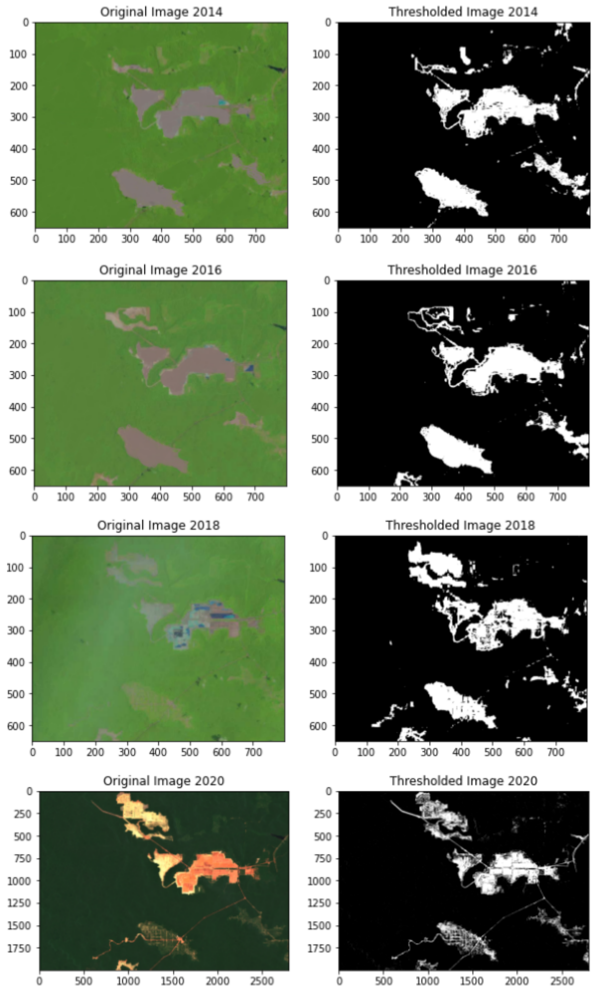


Figure 2.2.4 : Evolution of another forest in 8 years

We get:

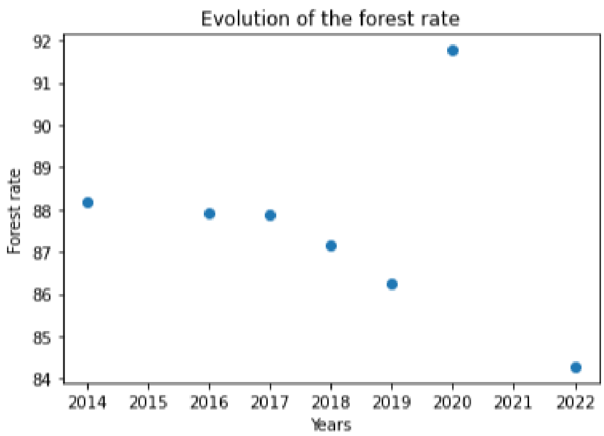


Figure 2.2.5 : Evolution of the forest rate

We can see that there is a sharper decline in recent years, with a regrowth in 2020, which we assume seems to be an error, due to a different value of the threshold.

We are going to apply a *gradient* in order to see more precisely what zones appeared in 4 years.

2.3 Application of a Gradient

To detect the differences between the images over several years, it is possible to use a time gradient. To do this, we try to see if the differences between two images is contrasted or not, and then apply a Taylor approximation to our gradient.

We first apply a pre-processing to the three images 2019 to remove the clouds. To do so, for each pixel, we keep the color (red, green or blue) which has the most intense value. For the year 2019, we have clouds above the forest. We will therefore put them all in green so that they can blend in with the trees.

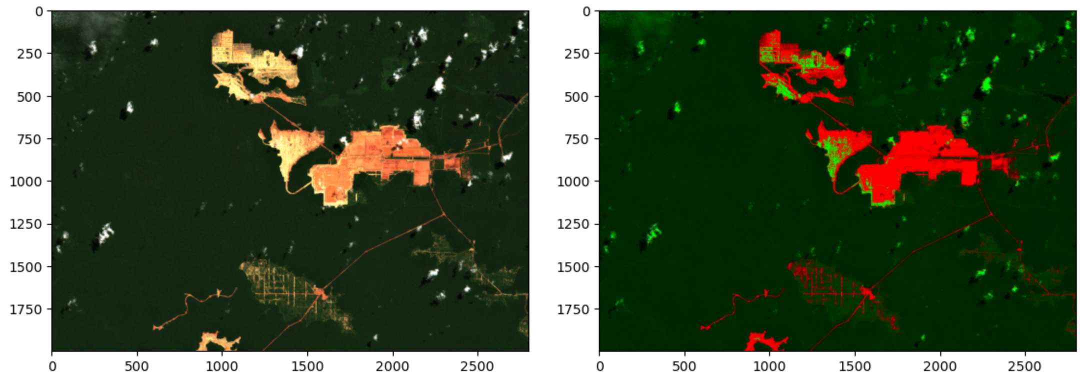


Figure 2.3.1: Pre-processing 2019 - removing clouds

We then set all the pixels to the same value using a threshold. Thus we separated the forest (in green) and the deforested area (in red).

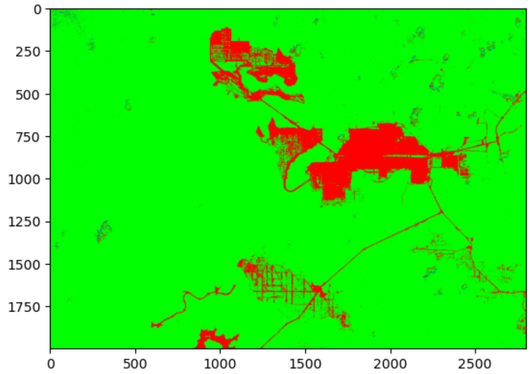


Figure 2.3.2: Pre-processing 2019

We do the same pre-processing for the images of 2020 and 2022 (without the need to remove the clouds), Then we apply a time gradient between years by approximating the gradient to the subtraction of the 2 images:

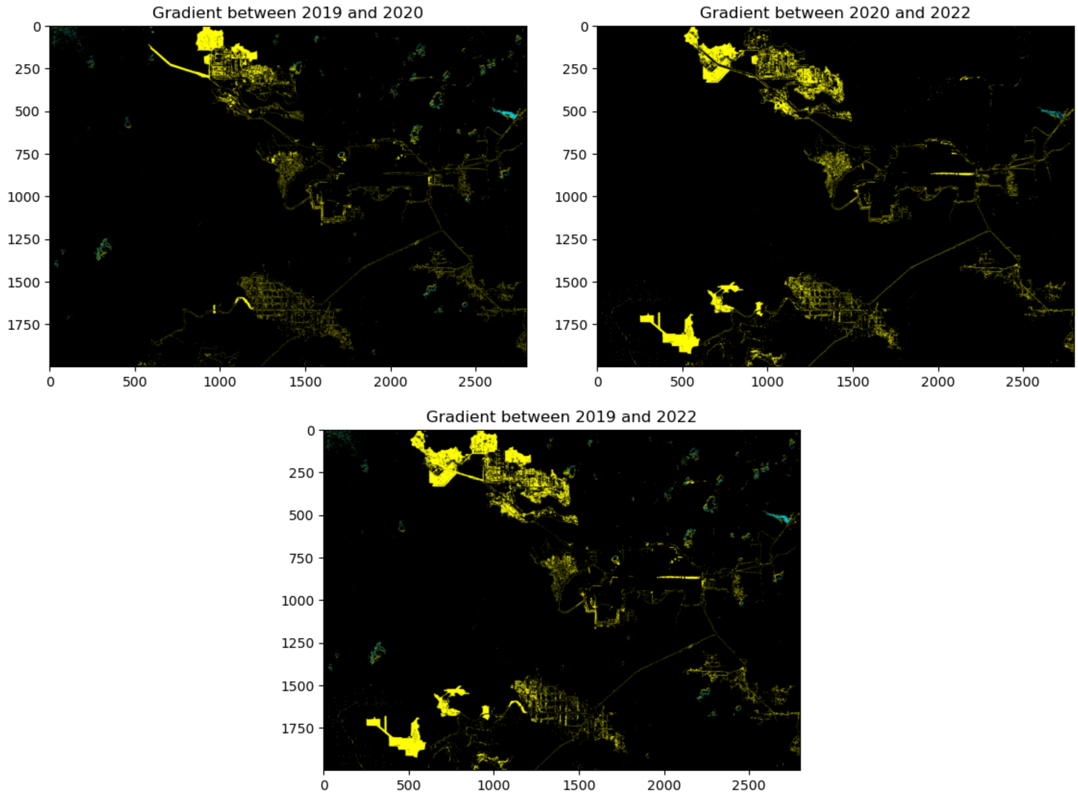


Figure 2.3.3: Application of a time gradient

The more the areas are yellow, the more the areas have changed. We can then count the number of pixels being yellow to deduce the increase in the rate of deforestation over the years.

Between 2019 and 2020, we obtain an increase in deforestation of $2,3\%$, between 2020 and 2022, we obtain an increase rate of $3,7\%$. Finally, between 2019 and 2022, this rate reaches $4,1\%$.

These results show distortions for the year 2020 where we had found that there had been a peak in reforestation, but here we find a general decrease. This is due to errors related to the different threshold values for the previous images. It is therefore seen that the gradient method gives a more precise result.

Here is an application of a time *gradient* on other zone where we had more data, get by *EarthExplorer*:

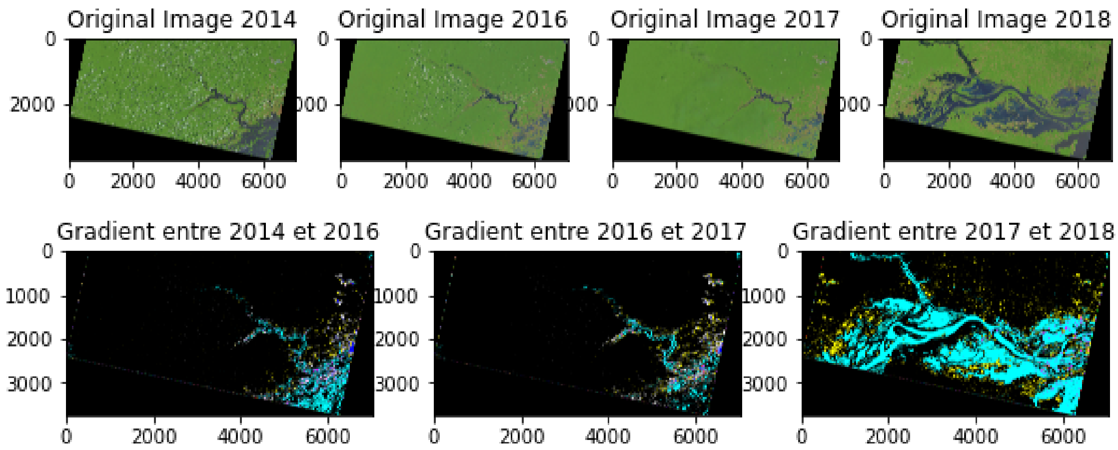


Figure 2.3.4 : Application of a time gradient on an other zone

We can clearly see here the evolution of deforestation over time, the blue area represents the area that has changed. Unfortunately, the clouds are often in the way, it would be necessary to study an algorithm to remove the clouds.

3 Second Approach - Exploitation Of The Chlorophyll Band

3.1 Description Of Chlorophyll

Chlorophyll is a characteristic pigment of green plants, located in cell organelles (chloroplasts) of the leaf parts of the plant and ensuring absorption of part of the sun's energy, which makes photosynthesis possible **Chlorophyll** [3].

Here is the absorption spectrum of the *Chlorophyll*:

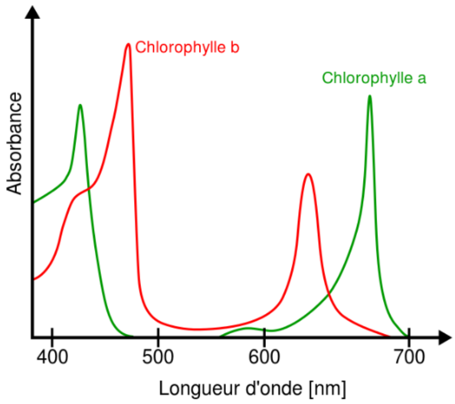


Figure 3.1.1 : Absorption spectrum of chlorophyll (Wikipedia)

Here is the absorption spectrum of the satelite *Sentinel 2*:

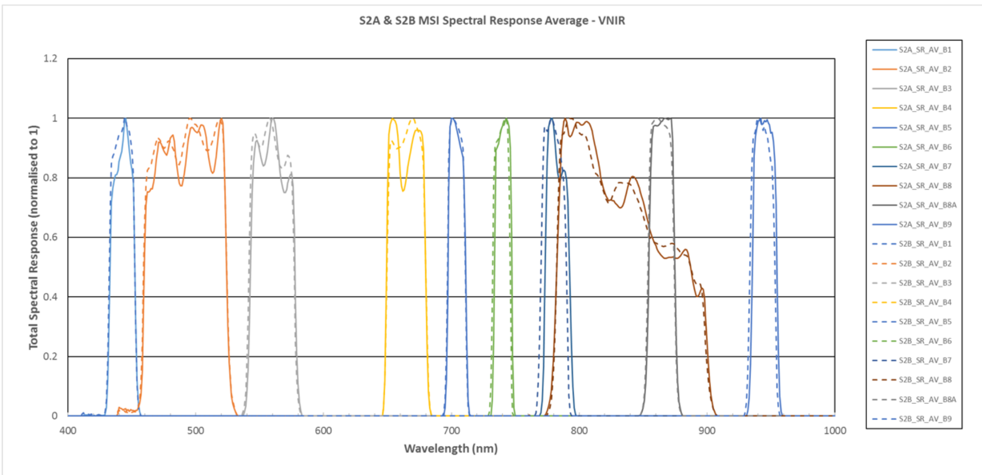


Figure 3.1.2 : Absorption spectrum of sentinel 2 (ESA)

This figure shows in dotted lines the values after updates. We can see that the *Sentinel 2* satellite has an absorption peak between $600nm$ and $700nm$ (*band 4*) which matches the peak of chlorophyll's absorption spectrum. If we reduce the tree search to a single band between $600nm$ and $700nm$, we increase the accuracy of our measurements.

3.2 Application of a Thresholding

We can thus apply a new threshold on these images:

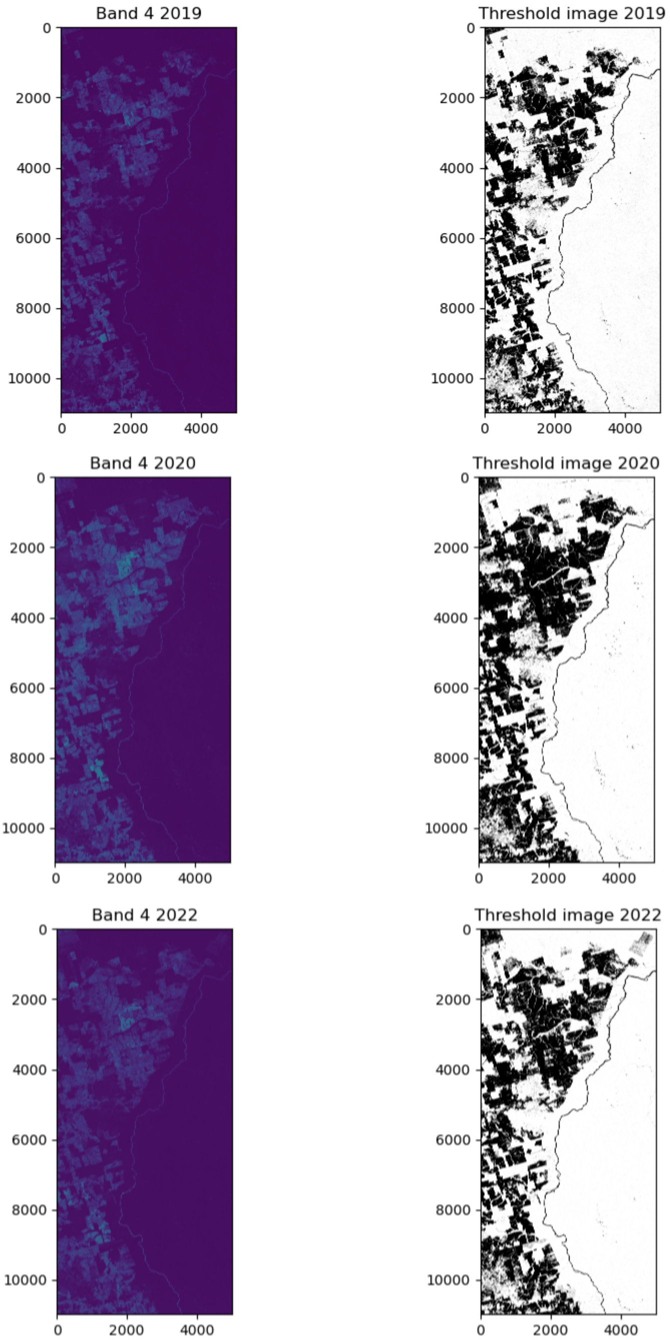


Figure 3.1.3 : Thresholded image of band 4

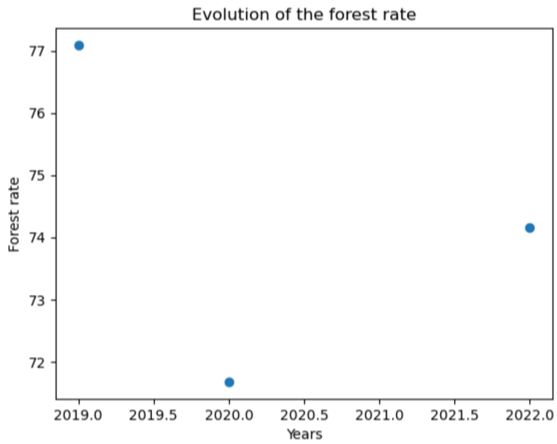


Figure 3.1.4 : Evolution of the forest rate

Here again, we have a significant peak in 2020 which seems to be due to the difference in thresholding that we had to applied to our images: each time, the image in 2020 had a different contrast which forced us to modify the threshold's value.

3.3 Application of a Gradient

We can also apply a *time gradient* as we did before:

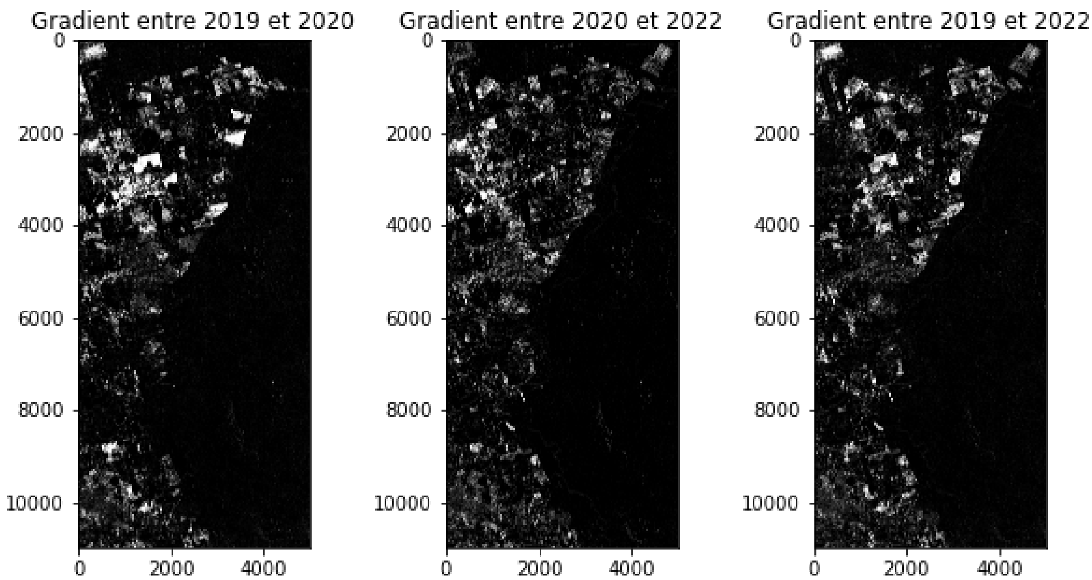


Figure 3.2.1 : Application of time gradient

We can then calculate each year the evolution of the deforestation gradient, And calculate the evolution of the deforestation rate with our first algorithm. **We deduce that between 2019 and 2020, there is a deforestation gradient of** 7.4%. **between 2020 and 2022, there is**

a deforestation gradient of 6.3% **and between 2019 and 2022 obtaining a deforestation gradient of** 7.2%**.**

We can therefore see that the gradient method is much more accurate but has always the problem of thresholding that appeared before: The threshold values in 2020 are always different from the others.

4 Conclusion

To conclude, our project focused on satellite-based detection of the Amazon rainforest to quantify and describe the evolution of deforestation. We have achieved promising results, with the most accurate outcomes obtained using the gradient method. However, we encountered challenges when applying thresholding techniques to a single RGB image due to varying contrasts, particularly in 2020.

Despite these difficulties, our research has provided valuable insights into the state of the Amazon rainforest and its ongoing deforestation. By averaging over the pixels for the threshold and developing a program to remove clouds regardless of the image used, we have enhanced the accuracy and reliability of our analysis. This approach has enabled us to effectively quantify and describe the forest's changes over time.

Looking ahead, our future objectives involve expanding our methodology to cover the entire Amazon region and eventually extending it to all forests worldwide. We aim to establish a comprehensive framework for forest quantification that incorporates innovative techniques and addresses challenges specific to different regions and forest types.

References

[1] Zero Deforestation : www.zero-deforestation.org

[2] Université Paris 1 Panthéon Sorbonne : www.e-cours.univ-paris1.fr

[3] Chlorophyll Wikipedia